



The Revised Child Anxiety and Depression Scale-Short Version: Scale reduction via exploratory bifactor modeling of the broad anxiety factor.

Citation

Ebesutani, Chad, Steven P. Reise, Bruce F. Chorpita, Chelsea Ale, Jennifer Regan, John Young, Charmaine Higa-McMillan, and John R. Weisz. 2012. "The Revised Child Anxiety and Depression Scale-Short Version: Scale Reduction via Exploratory Bifactor Modeling of the Broad Anxiety Factor." *Psychological Assessment* 24 (4): 833–845. doi:10.1037/a0027283.

Published Version

doi:10.1037/a0027283

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:34262172>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

RUNNING HEAD: RCADS short version

The Revised Child Anxiety and Depression Scale - short version: Scale reduction via exploratory bifactor modeling of the broad anxiety factor

Chad Ebesutani¹

Steven P. Reise²

Bruce F. Chorpita²

Chelsea Ale³

Jennifer Regan²

John Young⁴

Charmaine Higa-McMillan⁵

John R. Weisz⁶

Department of Psychology
*Yonsei University, Seoul, South Korea*¹

Psychology Department
*University of California at Los Angeles, Los Angeles, California*²

Department of Psychiatry and Human Behavior
*University of Mississippi Medical Center, Jackson, MS*³

Psychology Department
*University of Mississippi, University, Mississippi*⁴

Psychology Department
*University of Hawaii, Hilo, Hawaii*⁵

Psychology Department
*Harvard University and Judge Baker Children's Center, Boston, Massachusetts*⁶

Author Note

Chad Ebesutani, Department of Psychology, Yonsei University; Steven P. Reise, Psychology Department, University of California at Los Angeles; Bruce F. Chorpita, Psychology Department, University of California at Los Angeles; Chelsea Ale, Department of Psychiatry and Human Behavior, University of Mississippi Medical Center; Jennifer Regan, Psychology Department, University of California at Los Angeles; John Young, Psychology Department, University of Mississippi; Charmaine Higa-McMillan, Psychology Department, University of Hawaii at Hilo; John R. Weisz, Psychology Department, Harvard University and Judge Baker Children's Center.

Send correspondence to Chad Ebesutani, Department of Psychology, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul, South Korea 120-749. Email: ebesutani@yonsei.ac.kr

Abstract

Using a school-based (N=1060) and clinic-referred (N=303) youth sample, the present study sought to develop a 25-item shortened version of the Revised Child Anxiety and Depression Scale (RCADS) using Schmid-Leiman exploratory bifactor analysis to reduce client burden and administration time and thus improve the transportability characteristics of this youth anxiety and depression measure. Results revealed that all anxiety items primarily reflected a single “broad anxiety” dimension, which informed the subsequent development of a reduced 15-item Anxiety Total scale. Although specific *DSM*-oriented anxiety subscales were not included in this version, the items comprising the Anxiety Total scale were evenly pulled from the various anxiety-related content domains from the original RCADS (i.e., 3 items each from the generalized anxiety, separation anxiety, panic, social anxiety and obsessive-compulsive disorders subscales). The resultant 15-item Anxiety Total scale evidenced significant correspondence with anxiety diagnostic groups based on structured clinical interviews. The scores from the 10-item Depression Total scale (retained from the original version) were also associated with acceptable reliability in the clinic-referred and school-based samples ($\alpha = .80$ and $.79$, respectively); this is in contrast to the alternate 5-item shortened RCADS Depression Total scale previously developed by Muris, Meesters, and Schouten (2002) which evidenced depression scores of unacceptable reliability ($\alpha = .63$) in both present samples. The shortened RCADS developed in the present study thus balances efficiency, breadth and scale score reliability in a way that is potentially useful for repeated measurement in clinical settings as well as wide-scale screenings that assess anxiety and depressive problems. These future applications are discussed, as are recommendations for continued use of exploratory bifactor modeling in scale development.

Keywords: RCADS, Exploratory bifactor analysis, Measurement efficiency, Anxiety, Depression, Psychometric Properties

Approximately 8 to 27% of children and adolescents experience the debilitating effects of anxiety and depression at some point in their development (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003; Lewinsohn, Zinbarg, Seeley, Lewinsohn & Sack, 1997). Anxiety disorders in childhood have been related to increasing severity of anxiety symptoms, additional anxiety disorders, and depression in adolescence and adulthood (Bittner, et al., 2007). Due to the covert nature of anxiety and depression, children are often not recognized by teachers and parents as having emotional/behavioral problems until the symptoms cause significant interference in academic and social functioning (Muris & Meester, 2002). For this reason, self-report assessments may be particularly important to identify these problems in youth.

Numerous well-researched self-report instruments have been developed for the assessment of youth anxiety and depression. The list of available, scientifically-supported anxiety and depression measures includes the Children's Depression Inventory, the Multi-dimensional Anxiety Scale for Children, the Revised Children's Manifest Anxiety Scale, the Screen for Child Anxiety Related Emotional Disorders, the State-Trait Anxiety Inventory for Children, and the Youth Self Report. Although these measures assess a range of both dimensional and trait behaviors associated with anxiety or depression, they were not developed based on current Diagnostic and Statistical Manual (*DSM-IV*; APA, 2004) nosology, thereby questioning whether these measures assess the same constructs of pathological anxiety and depression. Perhaps due to this disconnect, some researchers have found limited discriminant validity with diagnostic categories for several of these most widely used child self-report measures of anxiety and depression (e.g., Stark & Laurent, 2001).

More recently, efforts have sought to supplement existing dimensional and trait measures of anxiety and depression with instruments that are more concordant with current *DSM* nosology

and provide scales that correspond more closely with specific dimensions of psychiatric disorders and thus may better inform diagnostic formulations (e.g., Birmaher et al. 1997). The Revised Child Anxiety and Depression Scale (RCADS) is one such measure, providing scales that index major depressive disorder (MDD) as well as the main features of five prominent *DSM-IV* anxiety disorders (separation anxiety disorder [SAD], social phobia[SP], generalized anxiety disorder [GAD], obsessive-compulsive disorder [OCD], panic disorder [PD]). Studies in both community and clinical samples have demonstrated strong psychometric properties for the RCADS as a favorable self-report measure of youth major depression and anxiety disorders (Chorpita et al., 2000, 2005). Unlike other anxiety measures (e.g., SCARED, RCMAS, MASC), the RCADS assesses both anxiety and depressive symptoms. This unique characteristic is clinically valuable given that children with anxiety problems are approximately 8.2 times as likely as children without anxiety problems to exhibit comorbid depression (Angold et al., 1999).

Although the five specific anxiety subscales is indeed a strength of the RCADS, its Anxiety Total Score (sum of all 37 anxiety items) and Depression (MDD) Total Score (sum of all 10 depression items) are arguably the most applicable RCADS scales with respect to identifying whether evidence-based treatments for depression and/or anxiety are needed. This is because *broad problem categories* (e.g., “anxiety problems”, as opposed to specific anxiety subtypes) are more concordant with the way treatment decisions are made currently in clinical practice, given the rare usage of structured diagnostic measures (Garland et al., 2003). Further, evidence-based treatments can also often be prescribed without precise diagnostic subtype determinations. That is, although some evidence-based treatments are in fact specific to certain anxiety subtypes (e.g., Coping Cat for separation anxiety disorder, generalized anxiety disorder, and social phobia, Kendall et al., 1997), much of the evidence base on treatment protocols has

demonstrated effectiveness with respect to youth characterized *not* by specific diagnoses, but rather, by syndrome elevations determined by cut-off scores on measures targeting broad problem areas (e.g., anxiety). This is reflected in randomized clinical trials (RCTs) testing the efficacy of treatment protocols. Specifically, only 41% of RCTs to date reported participant diagnoses among 435 randomized clinical trials reviewed by a recent study (Chorpita, Bernstein, & Daleiden, 2011). Therefore, from a treatment prescription perspective, the RCADS Depression Total Score and Anxiety Total Score are perhaps the most important and relevant scores to identify youth needing anxiety and depression treatment. Thus, a strength of the RCADS is the inclusion of both broad Anxiety and Depression Total scales, comprised of indicators from a variety of *DSM*-oriented specific content domains.

Instrument Characteristics Influencing Their Dissemination: Cost and Length

In clinical settings, empirically-based assessment tools, such as the RCADS, are important for aiding in accurate screening, informing treatment need and providing state-of-the-art care to patients. Unfortunately, in the business-driven reality of many clinical settings, the cost of empirically-supported assessment tools may be prohibitive. For example, in 2011, the RCMAS was advertised as starting at \$44.50, the YSR was \$50 for 50 forms and a scoring manual, the MASC was \$88 for a manual and 25 forms, the STAIC started at \$100 for 50 forms, and the CDI started at \$135 for the manual and 25 forms and scoring sheets. Community mental health clinics, non-profit organizations, and private offices that service large numbers of children with anxiety and depressive disorders may be unable to access these types of proprietary measures due to fiscal constraints. For this reason, some assessment developers have chosen to grant free access to their instruments. The 66-item SCARED-R, and the 47-item RCADS are

both available for free use¹. Largely due to their affordability and psychometric support, these measures are being used in many large-scale dissemination efforts. The RCADS, for example, has recently been instituted as a mandatory pre- and post-treatment measure for all youths whose primary targeted problem area is anxiety in the Los Angeles county mental health sector.

While cost and accessibility are highly relevant instrument characteristics, the actual length of the assessment is also particularly important to its transportability. In the past decade, there has been a shift toward the delivery of mental health services in school systems and in primary care settings in order to reach the most children and intervene as early as possible in the development of mental health disturbances (Rones & Hoagwood, 2000). This transition provides new opportunities to efficiently and effectively address children's mental health concerns, but these settings also bring about new demands for assessment measures and challenges for implementation. While instrument sensitivity and specificity are still vital, additional demands for instrument efficiency are just as important in community settings (Levitt, Saka, Romanelli & Hoagwood, 2007). Issues such as time required of youth to complete the measure, ease of scoring, and breadth of data collected are paramount to screening large numbers of children.

Balancing Reliability with Transportability

In 2002, Muris, Meesters, and Schouten recognized the need for a brief measure of anxiety and depression and endeavored to shorten the RCADS. Using standard factor analytic techniques, Muris and colleagues reduced the RCADS to 25 items. Based on a Dutch sample of 1748 school-based children ages 8-15 years old, factor analysis indicated that 25.5% of items loaded substantially onto a non-intended factor. Following confirmatory factor analysis, Muris and colleagues shortened the RCADS so that each subscale would contain just *five* items (via

¹ The 47-item RCADS measure and scoring program are available for free here:
<http://www.childfirst.ucla.edu/resources.html>

eliminating one GAD item, two SAD items, four social phobia items, four panic disorder items, and five MDD items). Muris and colleagues also summed the anxiety items to yield a 20-item Total Anxiety scale and the depression items to yield a 5-item Depression Total scale. Although this shortened RCADS version (Muris et al., 2002) has achieved certain goals of a briefer RCADS, additional test development efforts appear needed to ensure that the shortened RCADS meets the clinical demands of reduced test length, while maintaining adequate reliability of its broad scale scores without including redundant items or excluding potentially relevant items (e.g., OCD indicators). For instance, Muris and colleagues (2002) discarded all OCD items due to inconsistent loadings on the purported OCD factor in their school-based sample (i.e., only two OCD items loaded onto the purported OCD factor). However, given that eliminating the OCD items partially vitiates the goal of achieving wide content diversity in the assessment of anxiety symptoms, and previous studies have supported the RCADS OCD items as adequately assessing OCD symptoms across both US (Chorpita, et al., 2005; Chorpita, et al., 2000) and Australian (de Ross, Gullone, & Chorpita, 2002) samples, more research should be conducted before entirely eliminating the OCD scale and all of its items.

Additional attention also needs to be given to minimizing the inclusion of redundant items, such as *"I worry about bad things happening to me"* and *"I am worried that something bad will happen to myself"*, which are included on all RCADS versions to date. The inclusion of such synonymous items not only has the potential to confuse children due to redundancy, but also consumes valuable resources (e.g., time and test space) without providing incremental information to the assessment of youth symptoms. Further, although Muris and colleagues' (2002) shortened version nicely addresses scale length, it has recently been argued that scales comprised of so few items (such as their 5-item Depression Total scale) are likely associated

with problematic scale score reliability (see Emons et al., 2007). While developing briefer measures, it is thus important to find ways to shorten the entire measure (to achieve the goal of overall briefer assessment) without overly compromising the reliability of the scores from any given scale.

From a treatment identification perspective, as noted above, the RCADS Depression Total Score and Anxiety Total Score are arguably the two most important scales for which reliability should be maximized and prioritized in the scale reduction process. This perspective however was not taken while developing Muris and colleagues' (2002) version. An alternative test-reduction approach could therefore be applied to shorten the overall RCADS test length while better maintaining scale reliability of the Depression Total and Anxiety Total scale scores via (a) retaining the 10-item Depression Total scale (to not compromise its reliability), and, (b) reducing the 37-item Anxiety Total scale (as it comprises over three times as many items as the Depression Total scale). A shortened RCADS that maintains the original 10-item Depression Total scale and reduces the 37-item Anxiety Total score in a way that draws from the various anxiety content domains while excluding redundant items could better meet the demands of a shortened version. Such an approach may also employ exploratory bifactor analysis (see Reise et al., in press), given that this statistical technique has been shown to be particularly useful for identifying items (from multiple content domains) that load the strongest on the hypothesized general factor, such as "broad anxiety", to inform the development of the shortened Anxiety Total scale. No studies to date have applied this bifactor model to the RCADS anxiety items.

The Present Study

In the present study, we aimed to develop an alternate shortened version of the RCADS that prioritized shortening the 37-item Anxiety Total scale (to yield a briefer Anxiety Total scale)

and retained the original 10-item Depression Total scale. In this process, we first examined the factor structure of the Depression Total scale to ensure that its items adequately loaded on the purported Depression factor in both the school-based and clinic-referred samples. To select the items to comprise the reduced Anxiety Total score, we applied Schmid-Leiman bifactor exploratory factor analysis (EFA; Schmid & Leiman, 1957), which we viewed as a highly applicable factor analytic procedure to inform the selection of the “best” anxiety items from among the 37 RCADS anxiety items. Specifically, Schmid-Leiman EFA is an exploratory bifactor model (Holzinger & Swineford, 1939) which posits a common, general factor (e.g., “broad anxiety”) that reflects the common variance across all items (e.g., “broad anxiety” being a common underlying dimension related to all 37 RCADS anxiety indicators across the five anxiety content domains/subscales). Schmid-Leiman EFA also posits multiple orthogonal “group” factors (which explains any additional item response variance of each hypothesized content domain/scale above and beyond that accounted for by the common “broad anxiety” factor). In the present study, these “group” factors correspond to each anxiety subscale in that they represent specific content domains of the “broad anxiety” common factor.²

Taking this approach, we would be able to estimate the amount of variance in the “broad anxiety” common factor accounted for by each item (i.e., how well each item loaded on the “broad anxiety” factor), as well as how much variance was “left over” in each anxiety content domain (i.e., GAD, SAD, SP, PD, OCD) for each item after controlling for the “broad anxiety” factor. This approach would allow us to examine whether the OCD items adequately tapped the “broad anxiety” factor (supporting their inclusion in the Anxiety Total score – regardless of whether they adequately served as OCD indicators) and whether additional variance was left

² Although this analytic strategy was developed decades ago (Holzinger & Swineford, 1939), only recently has it re-emerged and been identified as a particularly useful approach to modeling the structure of certain types of psychopathology and various related outcomes (e.g., Gustaffson & Balke, 1993; Reise, Morizot, & Hays, 2007).

over in the OCD group factor after controlling for the “broad anxiety” factor supporting their specificity as OCD-related indicators. Given the mixed findings regarding the RCADS OCD factor across the examined US, Australian and Dutch samples (Chorpita, et al., 2000; de Ross, Gullone, & Chorpita, 2002; Muris et al., 2002), exploratory bifactor analysis appeared needed to inform which OCD items (if any) to retain, and more generally, which anxiety items should comprise the shortened Anxiety Total scale.

As outlined below, we used a cross-sample validation strategy to help inform which items to retain and to increase generalizability of the shortened measure. After identifying the items to comprise the shortened RCADS, we conducted validity analyses to examine the degree to which the reduced Anxiety Total Score and Depression (MDD) Total Score corresponded significantly with relevant (anxiety and depression) diagnostic groups using our clinic-referred sample.

Although we did not expect the shortened 15-item Anxiety Total scores to perform just as well as the original 37-item Anxiety Total scores due to its shorter length, we hypothesized that the shortened Anxiety Total scores would still correspond significantly with anxiety-related diagnostic groups. This hypothesis was based on our data analytic approach to include the anxiety items that loaded strongly on the “broad anxiety” common factor and tapped the various anxiety content domains.

Method

Participants

The school-based sample consisted of 1060 youth in grades 3-12 in public and private schools across the state of Hawaii (median grade = 7). Inclusion into the present study required youths being in grades 3-12 (cf. Chorpita et al., 2000) and having available RCADS data with no missing data. Of the 1215 RCADS forms with at least some data available, 1060 (87.2%) had

complete data with no missing values and were thus included in the present study. Demographic information for these 1060 school-based youth appears in Table 1.

The clinic-referred sample consisted of 303 children and adolescents referred from community agencies and public schools to two outpatient clinics to receive mental health assessments. Inclusionary criteria for this sample required youths being in grades 3-12 and having no missing RCADS data. Of the 332 RCADS forms with at least some data available, 303 (91.3%) had complete data with no missing values and were thus included in the present study. Youth ages ranged from 7 to 14 years ($M = 10.3$, $S.D. = 1.7$). Additional information regarding these youths' demographic information and diagnostic profiles appear in Table 1 and Table 2, respectively.

Procedure

The school sample was derived from a larger school-based study of negative emotions in youth. This study received Institutional Review Board approval at the University of Hawaii at Manoa. Parents provided signed consents on take-home forms that were returned to their children's school. All students provided assent prior to the administration and completion of questionnaires (including the RCADS) in a group format at school. Assistance was provided to students if they had difficulty reading and/or filling out questionnaires. All students were fluent in English, and each student received a \$5 gift certificate for participating.

Prior to any data collection from the clinic-referred sample, all youths and their legal guardians also underwent standardized Institutional Review Board-approved notice of privacy and consent procedures. Each youth was asked to complete a battery of questionnaires (including the RCADS) in English. Each youth was also administered a structured diagnostic clinical interview (i.e., the Children's Interview for Psychiatric Syndromes – Child Version; ChIPS-C;

Weller, Weller, Fristad, Rooney, & Schechter, 1999). Assessors consisted of Ph.D. level clinical child psychologists and senior doctoral students in clinical psychology. Detailed training on the use of the interview was provided to all assessors prior to administration with clients. This entailed didactic instruction on the use of the measure, observation of three ChIPS interviews conducted by trained assessors, and conducting a series of five ChIPS interviews while being observed by a trained assessor. Before interviewers were certified as proficient it was required that they match the trained assessor on all clinical diagnoses in three of their five training administrations, as well as Clinical Severity Ratings (CSRs) within at least one point on all assigned diagnoses. CSRs ranged from 0-10, with values ≥ 5 being indicative of clinically significant severity. The CSR procedures were based on commonly used ratings from a structured interview that is prevalently used in research (Silverman et al., 1996). Expert diagnosticians also provided supervision and feedback to assessors while formulating diagnoses.³

Measures

Revised Child Anxiety and Depression Scale (RCADS; Chorpita, Yim, Moffitt, Umemoto, & Francis, 2000) – The RCADS is a 47-item self-report questionnaire used to assess youth depression and anxiety symptoms consistent with *DSM-IV* nosology. The RCADS consists of five anxiety subscales corresponding to SAD, GAD, OCD, social phobia, and panic disorder. The 37 anxiety items may be summed to yield an Anxiety Total score. The RCADS also includes a major depressive disorder (MDD) scale comprising 10 items (also referred as the Depression Total score). Youth are asked to indicate how often each item applies to them according to a 4-point scale (0 = never, 1 = sometimes, 2 = often, 3 = always). Strong support for the validity and

³ A few diagnoses were assigned for problem areas not assessed formally by the ChIPS interview (e.g., trichotillomania). In instances whereby such problems were endorsed, the assessors obtained information from the youth directly regarding these problems according to *DSM* diagnostic criteria (APA, 2004). Supervisors and expert consultants were also available for consultation.

reliability of the RCADS scale scores have been demonstrated in both clinical (Chorpita, Moffitt, & Gray, 2005) and nonclinical samples (Chorpita et al., 2000; de Ross, Gullone, & Chorpita, 2002; Van Oort, Greaves-Lord, Verhulst, Ormel, & Huizink, 2009).

Children's Interview for Psychiatric Syndromes - Child Version (ChIPS-C; Weller, B., et al., 1999) The ChIPS is a semi-structured clinical interview designed to be administered to youths (ages 6-18 years) by trained interviewers. The ChIPS interview screens for 20 different *DSM-IV* Axis I disorders and psychosocial stressors. Content and concurrent validity, as well as inter-rater agreement of the scores derived from the ChIPS have been demonstrated in previous studies in clinical and community samples (e.g., Fristad et al., 1998a,b; Teare et al., 1998a,b).

Data Analytic Approach

Our data analytic approach to shortening the RCADS was as follows. We first employed Schmid-Leiman bifactor exploratory factor analysis (EFA) among the 37 anxiety items, positing a “broad anxiety” factor common to all anxiety items, and five group factors corresponding to the five hypothesized anxiety subdomains (see Reise, Moore, & Haviland, 2010, for a more detailed description of the Schmid-Leiman bifactor EFA procedures). We conducted these analyses using oblique (promax) rotation given that orthogonalization under the Schmid-Leiman procedures (Schmid & Leiman, 1957) involves first estimating a correlated traits model using an oblique factor rotation (see Reise, Moore, & Haviland, 2010). Due to the RCADS response choices being categorical (ordinal), we used polychoric correlation matrices (Holgado-Tello, Chacón-Moscó, Barbero-García & Vila-Abad, 2010) and the minimum residual (OLS) solution estimation method (Harman, & Jones, 1966), available in the PSYCH package in the R statistical software (R Development Core Team, 2008).

We also computed coefficient omega hierarchical (Zinbarg et al., 1997; 2005) which is a statistic based on the bifactor model that estimates the proportion of variance in raw scores attributable to a single general trait (e.g., “broad anxiety”). When calculating this statistic, variation in scores due to group factors are treated as nuisance variance and thus relegated to measurement error. This statistic allows one to examine the extent to which scores from a measure reflect a single dimension, such as “broad anxiety” common to all 37 RCADS anxiety items (see Gustafsson & Aberg-Bengtsson, 2010). We calculated this statistic in order to examine the degree to which the resultant shortened Anxiety Total scale score represents multidimensional constructs of specific anxiety subdomains, or primarily a single, overarching anxiety factor. Coefficient omega hierarchical may also be used to obtain a less biased estimate of the reliability of scale scores – the difference between the coefficients alpha and omega hierarchical indicates the degree to which the reliability estimate is affected by group factors biasing the estimate of true score variation (see Reise et al., 2010).

As noted above, our data analytic strategy was also based on a cross-sample validation strategy. We conducted separate Schmid-Leiman bifactor EFAs on the clinical and school-based samples, and we examined both sets of the results to identify three anxiety items per anxiety content domain that met the following criteria across both samples: (a) loaded strongly (i.e., $>.30$; McDonald, 1999) on the common “broad anxiety” factor (to include the “best” anxiety items among the 37 anxiety items that most strongly measured the “broad anxiety” factor and to justify summing the included anxiety items to yield the Anxiety Total score), (b) loaded strongly (i.e., $>.30$) on the hypothesized specific content domain (supporting each item as a strongly-representative indicator of its associated anxiety domain to increase content diversity), (c) loaded

“simply” (to identify items that *uniquely* tapped the corresponding content domain), and (d) were not synonymous (re-worded) items (to further increase content diversity).⁴

Regarding the (MDD) Depression Total scale, we used standard EFA procedures to examine how well the MDD items loaded on the hypothesized depression factor. We did not subject the MDD subscale to Schmid-Leiman factor analysis (as done above with the anxiety items) given that the RCADS MDD scale has historically been demonstrated to be a unidimensional construct without subfactors (Chorpita et al., 2000); other researchers have also recently demonstrated that the construct of depression as measured by diverse instruments does not appear to be a multidimensional construct (Aggenn, 2005). As with the previous bifactor EFAs, we used oblique (promax) rotation and polychoric correlation matrices with the minimum residual (OLS) solution estimation method for these standard EFAs. We reported all factor loadings for each item on the hypothesized depression factor, and we retained all depression items if they loaded significantly on the depression factor (i.e. $> .30$).⁵ We also examine scale score reliability of the 10-item and 5-item Depression Total scale given noted concerns of the reliability of Muris and colleagues' (2002) shortened 5-item Depression Total scale. We used .70 as the cut-off for adequate reliability (Nunnally, 1978). We could not compute coefficient omega hierarchical given that this statistic applies only to bifactor models.

Concurrent Validity with Clinical Diagnostic Groups

⁴ Although we could have selected only two items from each of the various anxiety content domains to comprise the reduced Anxiety Total scale to achieve an even shorter Anxiety Total scale, we decided to select three items per content domain given that: (a) three items just identifies a factor, thereby allowing future factor analytic research to be conducted on the purported anxiety subscales, and relatedly, (b) including only two items per content domain would preclude such research from being conducted. To reiterate, our intention underlying this decision was not to develop interpretable subscales, as seen in Muris and colleagues' (2002) shortened version; instead this decision allows future researchers to continue studying the factor structure of the purported anxiety subscales/content domains relative to the general anxiety factor. This is particularly important given the "broad anxiety" framework being proposed in the present study as representing the structure of and relationship between the various anxiety items across the anxiety content domains.

⁵ Although we used .30 at the cut-off to indicate a salient and meaningful loading across the various EFA analyses, we used .30 as a general guide for interpretation as opposed to as a rigid cut-point due to the lack of empirical data supporting this value as a strict factor loading cut-off.

To examine the concurrent validity of the shortened RCADS scores with external (and independent) criteria, we used the clinic-referred sample to examine the correspondence of the shortened 15-item Anxiety Total scale with relevant anxiety-related diagnostic groups. To do this, we conducted ANOVAs and calculated receiver operating characteristic (ROC) area under the curve (AUC) values using Analyze-It for Microsoft Excel version 2.12 (Analyze-It Software Ltd., 2008). An ROC AUC value provides an approximation of the degree to which a scale (e.g., the reduced Anxiety Total scale) accurately predicts the presence or absence of a diagnostic category, such as anxiety problems (cf. Ferdinand, 2008). Larger AUC values represent better prediction power with respect to diagnostic group status, and AUC values significantly greater than .50 indicate that the scale can predict classification status better than chance. AUC values may also be interpreted according to the following (Hosmer & Lemeshow 2000): .50–.70 = “poor”; .70–.80 = “acceptable”; .80–.90 = “excellent”; .90–1.00 = “outstanding” discrimination.

We also compared the AUC values of the reduced Anxiety Total scale (developed in the present study) to the original 37-item Anxiety Total scale and to Muris and colleagues’ (2002) 20-item version (with respect to accurately predicting pertinent diagnostic groups) via *z*-test comparisons of AUC values (DeLong et al., 1988). We used the significance level of $p < .01$ to control for Type-I error rates given the number of tests conducted (i.e., $.05/4 = .01$).

Results

Scale Development

Anxiety Total Scale. Results of the Schmid-Leiman EFA (based on the 37 anxiety items) appear in Table 3. Results demonstrated that all anxiety items loaded adequately ($>.30$) on the “broad anxiety” factor in both samples (see Table 3, columns “A”). These results support the notion of there being a “broad anxiety” factor common to all the anxiety items, and that all

anxiety items adequately tapped this general factor. Relatedly, coefficient omega hierarchical was equal to .74 and .71 in the clinic-referred and school-based samples, respectively. This suggests that 74% and 71% of the variance of the Anxiety Total composite scores could be attributed to variance on the “broad anxiety” factor, which further supports that the scores obtained from the 37 anxiety items are largely reflective of a single common source (i.e., “broad anxiety”) and thus need not be scored using multidimensional scoring procedures, but as a primarily unidimensional anxiety construct.⁶

Further, each of the five anxiety content domains had at least three items that loaded adequately on the hypothesized content domain and simply on no other subfactor. The items that loaded adequately and simply on the hypothesized content domain across *both* the clinical and school-based samples appear in bold in Table 3. These items were considered to be among the best items for potential inclusion in the shortened Anxiety Total scale. There were two rationally-determined exceptions to this rule. First, although the GAD item #37 (“*I think about death*”) did not load on any content domain based on the *school* sample, this item did load adequately ($>.30$) on the “broad anxiety” factor in both samples, as well as adequately and simply on the hypothesized GAD content domain in the *clinical* sample. Therefore, despite not loading on the content domain in the school sample, we decided to retain this item for further examination given that all other criteria were met and this item may be particularly clinically useful (e.g., related to potential suicidal ideation). Second, although no OCD items loaded on any content domain based on the *school* sample (after controlling for the “broad anxiety” factor), we

⁶ Notably, alpha coefficients for the 37-item Anxiety Total scale based on the clinic-referred and school-based samples were .96 and .94, respectively. These results suggest that these alpha coefficient-based reliability estimates of the Anxiety Total score represent inflated reliability estimates due to group factors biasing the estimate of true score variation. Local dependence can also spuriously increase estimates of reliability (e.g., Wainer, & Thissen; 1996), which is another problem with computing and interpreting alpha coefficient reliability estimates. As argued elsewhere (e.g., Reise, Moore, & Haviland, 2010), coefficient omega hierarchical reflects a more accurate estimate of the reliability for composites such as the Anxiety Total score (which comprises multiple group factors), and should be interpreted as such instead of coefficient alpha to obtain a more accurate reliability estimate.

decided to retain three OCD items (i.e., RCADS #31, 42, 44) given that those three items loaded simply on the hypothesized OCD content domain in the *clinical* sample, as well as loaded adequately on the “broad anxiety” factor in both samples. Further, retaining the OCD items may further increase the content diversity of the Anxiety Total scale and may also be particularly useful when assessing clinical samples with arguably higher base rates of OCD symptoms.

Separation anxiety disorder items. As seen in Table 3, there were four (bolded) SAD items from which three items were to be selected to represent this content domain. We decided to discard RCADS #46 and retain the remaining items (RCADS #5, 17, 33). Although RCADS#46 loaded the highest on the SAD content domain in the school sample (.62), it loaded the lowest on the SAD content domain in the clinical sample (.32); moreover, this item #46 (“*I would feel scared if I had to stay away from home overnight*”) has the potential to be less applicable to younger youth (due to younger children typically not staying away from home overnight). The other three items loaded strongly on the “broad anxiety” factor in both sample and were also content diverse within the area of separation anxiety problems.

Generalized anxiety disorder items. As seen in Table 3, there were four (bolded) GAD items from which three items were to be selected to represent this domain. We decided to discard RCADS #22 (“*I worry that bad things will happen to me*”) due to being too synonymous (and thus redundant) with RCADS #27 (“*I worry that something bad will happen to me*”), which loaded the highest on the “broad anxiety” common factor in both the clinical (.70) and school (.71) samples. We retained the remaining three RCADS items (#13, 27 and 37) which loaded strongly on the “broad anxiety” factor.

Panic disorder items. As seen in Table 3, there were five (bolded) panic disorder items from which three items were to be selected to represent this domain. We included RCADS #41

(*“I worry that I will suddenly get a scared feeling when there is nothing to be afraid of”*) and RCADS #26 (*“I suddenly start to tremble or shake when there is no reason for this”*) due to these items loading the highest on the “broad anxiety” factor in both the clinical (.70, .61, respectively) and school (.55, .51, respectively) samples. We excluded RCADS #39 (*“My heart suddenly starts to beat too quickly for no reason”*) due to the experience of a racing heart being a ubiquitous phenomenon. We thus retained RCADS #36 (*“I suddenly become dizzy or faint when there is no reason for this”*) as this was the next best item in terms of loadings on the “broad anxiety” factor and having a specific association with this content domain.

Social anxiety disorder items. As seen in Table 3, there were seven (bolded) social anxiety items from which three items were to be selected to represent this domain. Interestingly, the social anxiety items split into two related factors (*“perfectionism”* and *“social humiliation”*) in the school sample.⁷ We thus sought to select at least one item from each of these two sub-domains related to social anxiety. We included RCADS #4 (*“I worry when I think I have done poorly at something”*) given that this item loaded the highest on the “perfectionism” content domain in both the clinical (.51) and school (.50) samples, as well as adequately on the “broad anxiety” factor in both samples. We also included RCADS #32 (*“I worry what other people think of me”*) and RCADS #43 (*“I feel afraid that I will make a fool of myself in front of people”*) given that both items contain an audience related to social anxiety (i.e., *“people”*), and these were the two items that loaded the highest on the “broad anxiety” factor in both the clinical (.55, .55, respectively) and school (.48, .50, respectively) samples.

⁷ We confirmed that the social anxiety items split into these two social anxiety-related factors specifically in the school sample by conducting standard EFA procedures on only the nine social anxiety items using both the clinical and school-based samples. We examined 1-factor, 2-factor and 3-factor EFA solutions of the nine social anxiety items. Results revealed that the social anxiety items formed a single factor in the clinical sample (loadings ranged from .65 - .76) and split into the “perfectionism” factor (loadings ranged from .40 - .73) and “social humiliation” factor (loadings ranged from .37 to .85) in the school sample.

Obsessive-compulsive disorder items. As seen in Table 3, there were only three (bolded) OCD items that loaded on the OCD factor in the clinical sample. We thus retained these three items to represent this domain. To ensure that these arguably low base-rate OCD-related symptoms were adequately endorsed in both the school and clinical samples, we examined frequencies of each of their response choices. All response options (i.e., “never”, “sometimes”, “often”, “always”) were endorsed with adequate and reasonable frequencies (ranging from 4% to 67% in the clinical sample and 6% to 55% in the school sample), supporting the inclusion of these items. In addition to loading simply on the OCD factor in the clinical sample, all three OCD items also loaded strongly on the “broad anxiety” factor in both samples.

This process led to the selection of 15 items to comprise the shortened Anxiety Total scale (three items from each of the five content domains). We then re-conducted Schmid-Leiman bifactor EFA to confirm the structure and loadings of these 15 items. Across both samples, as expected, the 15 items loaded highly on the “broad anxiety” factor and simply on the expected content domains in both samples (results available from the first author). Coefficients alpha and omega hierarchical for these 15 anxiety items in the school sample were .86 and .71, respectively, and .91 and .76 in the clinical sample, respectively, again demonstrating coefficient alpha as an inflated reliability estimate relative to coefficient omega hierarchical. It is important to note, however, that by standard conventions of interpretation for both indices, the shortened scale exhibited adequate reliable psychometric properties (Nunnally, 1978).

Depression Total Scale. As expected, all depression items loaded highly on the depression factor in the clinical sample (ranging from .49 to .69) and school sample (ranging from .52 to .73). Two- and three-factor solutions were not interpretable. Alpha coefficients for the 10-item Depression scale were .80 and .79 in the clinical and school sample, respectively.

Alpha coefficients for Muris and colleagues' (2002) reduced 5-item Depression Total scale were substantially lower ($\alpha=.63$ in both samples) and fell below the cut off for acceptable reliability. These results supported retaining all 10 MDD items to comprise the Depression Total scale.

Scale Validation

RCADS Anxiety Total scale. As predicted, the original 37-item Anxiety Total scale, 20-item shortened Anxiety Total scale (developed by Muris and colleagues, 2002) and the 15-item Anxiety Total scale (developed in the present study) discriminated youths with any anxiety disorder diagnosis ($n=82$) from youths without any anxiety diagnosis ($n=217$), as evidenced by significant between-group ANOVAs ($p < .01$) and AUC values significantly greater than .50 ($p < .01$). Although the AUC value of the 15-item shortened Anxiety Total scale developed in the present study was the greatest ($AUC=.74$, $SE=.03$), it did not differ significantly from the AUC of the original Anxiety Total scale ($AUC=.72$, $SE=.04$; $z=1.30$, $p=.19$) or from Muris and colleagues' (2002) Anxiety Total scale ($AUC=.71$, $SE=.04$; $z=1.77$, $p=.08$). The AUC of the original Anxiety Total scale and Muris and colleagues' (2002) 20-item Anxiety Total scale also did not significantly differ, $z = 0.91$, $p = .36$. All AUC values fell in the "acceptable" classification category.

RCADS MDD (Depression Total) Scale. As predicted, the 10-item Depression Total scale and Muris and colleagues' (2002) 5-item Depression Total scale both discriminated youths with any affective disorder (MDD, dysthymic disorder, depressive disorder not otherwise specified; $n=36$) from youths without any affective disorder ($n=263$), $F=17.12$, $p < .01$; $F=19.12$, $p < .01$, respectively. AUC values were significantly greater than .50 ($p < .01$), and the AUC value of the original 10-item Depression Total scale ($AUC=.73$; $SE=.04$) did not differ from the AUC value

of Muris and colleagues' (2002) 5-item Depression Total scale ($AUC=.73$; $SE=.04$), $z=0.00$, $p=1.00$. Both AUCs fell in the “acceptable” classification category.

Normative Data

Lastly, to allow for T-scores to be calculated for each of the RCADS short version scales, we have provided normative data in Table 4 based on the current school sample ($N=1060$), including means and standard deviations for the RCADS short version MDD scale, Anxiety Total scale, and RCADS Total scale — broken down by gender and grade level.

Discussion

Using Schmid-Leiman exploratory bifactor analysis, we shortened the RCADS from 47 items to 25 items using a large school-based sample and a clinic-referred sample of youth. Although a previous attempt was made to shorten the RCADS that also led to a 25-item RCADS comprising 5 items per scale (Muris et al., 2002), we took a substantially different approach to developing a shortened RCADS. First, our item selection procedures relied primarily on the unique Schmid-Leiman exploratory bifactor analytic procedure that posits a general factor (presumed to underlie all anxiety items) and estimates the degree to which each item loads on this general factor as well as on specific content domains. More specifically, we retained three items per anxiety content domain that loaded highly on the “broad anxiety” factor, simply on the hypothesized domain and were not redundant items. This is the first study of the RCADS that took an exploratory bifactor analytic approach to examine the degree to which each anxiety indicator represents the “broad anxiety” factor. The exploratory bifactor model fit the data well, demonstrating that the variability in scores derived from the anxiety items – although representative of their respective content domains – are largely accounted for by this “broad anxiety” factor common to all anxiety items and content domains.

In our scale reduction process, we were also the first to minimize content overlap by eliminating synonymous items, such as eliminating “*I worry that bad things will happen to me*” (RCADS #22) due to this item being a re-worded version of RCADS #27 (“*I worry that something bad will happen to me*”). Notably, both items were included in Muris and colleagues’ (2002) reduced scale, demonstrating one of the ways in which our approach was unique with respect to item reduction. It is also notable that we decided to not reduce the original 10-item Depression Total scale, which differed in approach relative to Muris and colleagues’ (2002) shortened version. We made this decision given that having too few items per scale can degrade reliability to unacceptable levels (Emons et al., 2007). Our results supported this view; although AUC values of the depression scales did not differ in the concurrent validity analyses, the 10-item MDD scale was associated with acceptable alpha reliability estimates, whereas Muris and colleagues’ (2002) reduced 5-item MDD scale was associated with unacceptable alpha reliability estimates. While more studies should be conducted to further examine the reliability of both MDD scales – including test-retest analyses – the present finding support the approach taken in the present study to retain the original 10-items Depression Total scale (for the sake of reliability) and focus on reducing the length of the original 37-item Anxiety Total scale.

In this process of shortening the Anxiety Total scale, we decided to forgo the development of specific anxiety subscales for several reason, including research showing that assessing the broad dimension of general anxiety problems (as opposed to specific anxiety subtypes) is sufficient to inform the prescription of evidence-based treatments for anxiety (e.g., exposure treatment; Chorpita et al. in press) and developing anxiety *subscales* to include a sufficient number of items to achieve acceptable levels of reliability would make scale reduction unfeasible. Researchers have also recently argued that *total scores* should be interpreted instead

of specific subscales in many cases due to items often serving as better indicators of the general factor (e.g., “broad anxiety”) rather than of specific subdomains (e.g., Reise, et al, 2010). Our results supported this idea, given that over 70% of the variance of the Anxiety Total score composite could be attributed to variance on the common “broad anxiety” factor. The RCADS anxiety items seem to be primarily measuring a single dimension (i.e., “broad anxiety”) albeit with some degree of content specificity (as evidenced by simple and adequate loadings on the anxiety content domains after controlling for the “broad anxiety” factor).

The issue surrounding retaining or discarding the OCD scale/items was also addressed in the present study. Interestingly, results based on the present *school sample* of youth were consistent with Muris and colleagues’ (2002) finding that the OCD items did not load on the specific OCD content domain. This finding, however, is in direct contrast to the outcome of the same process in the clinically-referred sample, wherein three OCD items did load simply on the hypothesized OCD factor. In addition, exploratory bifactor analysis revealed that all OCD items loaded adequately on the “broad anxiety” factor in both samples (supporting the inclusion of the OCD items in the Anxiety Total composite score). These results support retaining the three noted OCD items (particularly for use among clinical samples), and demonstrate the importance of conducting EFA studies on both school and clinical samples before discarding an entire scale as results may differ across clinical and non-clinical samples.

Interestingly, although both the present and Muris and colleagues’ (2002) shortened versions of the RCADS are 25 items long, there are only 13 overlapping items; Muris and colleagues’ (2002) Anxiety Total scale has 12 items not included in the present shortened Anxiety Total scale, and our shortened Anxiety Total scale has 12 items not included in their scale. This highlights the unique information provided by Schmid-Leiman exploratory bifactor

analysis and its ability to inform scale development. Examination of the degree to which each item accounts for variability in the general factor (i.e., “broad anxiety”) as well as each content domain represents a substantially different process and outcome from traditional factor analytic techniques, as evidenced by comparing the current solution to that of Muris et al. (2002).

Although Schmid-Leiman exploratory bifactor analysis has been seldom utilized in test construction studies over the past several decades (due to this technique not being available in most statistic programs and only recently re-emerging in the field; Reise et al., in press; Wolff & Preising, 2005), the present study demonstrated its appropriateness and that the Schmid-Leiman bifactor EFA approach may be well-suited for modeling data that are believed to have a single, common cause (e.g., “broad anxiety”) as well as associated with specific group factors. Future studies that seek to develop and/or shorten questionnaires comprising multiple subscales that may reflect a single common dimension should consider applying Schmid-Leiman exploratory bifactor analysis to examine each item’s contribution to the general factor as well as specific content domains. The Schmid-Leiman exploratory bifactor analytic technique is easily executed⁸ and available for free in the R statistical program (*psych* package; Revelle, 2009).

It is also important to note that the present study demonstrated that all Anxiety Total scale versions (i.e., the original 47-item version, Muris and colleagues’ (2002) 20-item version, and the present 15 item version) corresponded significantly with anxiety-related diagnostic groups. All scales also performed with “adequate” classification accuracy relative to child diagnoses derived from a clinical structured interview. Given these findings, the decision

⁸ With the *psych* package loaded in R, Schmid-Leiman exploratory bifactor analysis (positing a general factor and five content domains, as in the present study) may be executed on a polychoric correlation matrix (named “polychor_x” in this syntax example), using *oblimin* rotation and the *minimum residual (OLS)* estimation method with the following R syntax: `schmid(polychor_x, nfactors=5, fm="minres", digits=2, rotate="oblimin")`. A polychoric correlation matrix (“polychor_x”) first needs to be calculated based on raw data matrix “x” with the following R syntax: `polychor_x <- polychoric(x, polycor=TRUE, ML=FALSE, std.err=FALSE)`.

regarding which version to use may depend on the intended purpose(s) of assessment. For example, if information regarding specific anxiety subscales is desired, the original 47-item RCADS or Muris and colleagues' (2002) RCADS shortened version (which yields 5-item anxiety subscales), may be preferable due to provision of the *DSM*-oriented subscales. However, given that Muris and colleagues' (2002) RCADS shortened version does not include an OCD subscale and each subscale comprises only 5 items (including the Depression Total scale), the original RCADS may be preferred in contexts where more reliable and/or diversity of specific anxiety subscale scores are sought. On the other hand, in contexts where diagnostic specificity is less focal (e.g., the assessment of specific anxiety subscales is not needed) and the assessment goals include obtaining estimates of general anxiety and depressive problems via briefer assessment that maintains adequate reliability, the shortened 25-item RCADS developed in the present study may be preferable. For example, screening a large student body and providing general feedback to school administrations may not entail the need for detailed anxiety subscale scores. In these settings, measurement efficiency is often of paramount concern, given time concerns of assessing students during school hours and issues of developing and maintaining cordial, satisfying relationships with school systems (Chorpita & Mueller, 2008).

Despite the strengths of the current study, there were limitations that warrant attention. First, although the clinic-referred sample comprised youth from multiple regions in the US (including Hawaii and Massachusetts), the school sample constituted only school children living in Hawaii. Although this school-based Hawaii sample was ethnically diverse, the inclusion of students from other US regions may have further increased the generalizability of the present findings. Second, the utility of the shortened Anxiety Total score could have been further examined if there were a greater number of anxious youths in the clinical sample. For example,

having a larger sample of anxious youth may have allowed us to test, for example, the ability of the Anxiety Total score to discriminate between youths with anxiety (and no depression) from youths with depression (and no anxiety). Relatedly, future studies should also compare the ability of the various versions of the Anxiety Total scale to identifying youths with each type of anxiety disorder in terms of sensitivity and specificity. Such studies, however, would require much larger clinical samples that include youths with “pure anxiety” diagnostic profiles (e.g., social anxiety and no other anxiety). It is also notable that there are no independently-sampled normative data available for the RCADS shortened version to aid in its interpretability. Future studies should seek to collect normative data to increase the clinical utility of its scale scores. Lastly, there has been some recent work conducted that has demonstrated the feasibility and utility of developing computerized adaptive testing procedures to assess mental health problems such as anxiety and depression. Gibbons and colleagues (2010), for example, recently used item response theory and developed computerized adaptive testing procedures to measure mood and anxiety problems among adults. Instead of using a short fixed-length test — such as the shortened RCADS developed in the present study — these researchers were able to assess individuals' levels of anxiety and depression with a small bank of items pulled from a larger item bank of 616 items tailored to each individual's particular response patterns. Similar avenues of research could be pursued with the RCADS to develop an alternate adaptive testing method of assessing youth anxiety and depressive problems with this measure.

Despite these limitations and areas for future research, the present findings support the use of the 25-item RCADS for a more efficient assessment of the general problem areas of anxiety and depression. In addition to demonstrating the usefulness of Schmid-Leiman exploratory bifactor analytic technique in identifying items that best tap the common dimension

of “broad anxiety”, the present study also provides the field with a unique tool that addresses the intersection of evidence-based assessment strategies and “real-world” demands. We have also made this instrument available for free, as we hope that it may assist in achieving more wide scale assessment of youth anxiety and depressive problems in settings that face both fiscal and time constraints.⁹

⁹ The 25-item shortened version may be found in the Appendix or downloaded for free at the following url: www.childfirst.ucla.edu/resources.html

References

- Aggen, S., Neale, M., & Kendler, K. (2004). *DSM* criteria for major depression: evaluating symptom patterns using latent-trait item response models. *Psychological Medicine*, 35, 475–487.
- Angold, A., Costello, E. J. & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology & Psychiatry*, 40, 57-87.
- Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, 107, 238-246.
- Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J., & McKenzie Neer, S. (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36, 545–553.
- Bittner, A., Egger, H., Erkanli, A., Costello, E., Foley, D. & Angold, A. (2007). What do childhood anxiety disorders predict? *Journal of Child Psychology and Psychiatry*, 48, 1174-1183.
- Browne, M., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-159). Newbury Park, CA: Sage.
- Chorpita, B. F., Bernstein, A., & Daleiden E. L. (2011). Empirically Guided Coordination of Multiple Evidence-Based Treatments: An Illustration of Relevance Mapping in Children's Mental Health Services. *Journal of Consulting and Clinical Psychology*, 79(4), 470-480

- Chorpita, B. F., Daleiden, E., Ebesutani, C., Young, J., Becker, K., Nakamura, B., Phillips, L., et al. (in press). Evidence Based Treatments for Children and Adolescents: An Updated Review of Indicators of Efficacy and Effectiveness. *Clinical Psychology: Science and Practice*.
- Chorpita, B. F., & Mueller, C. W. (2008). Toward new models for research, community, and consumer partnerships: Some guiding principles and an illustration. *Clinical Psychology: Science and Practice*, 15, 144-148.
- Chorpita, B. F., Moffitt, C., & Gray, J. (2005). Psychometric properties of the Revised Child Anxiety and Depression Scale in a clinical sample. *Behaviour Research and Therapy*, 43, 309-322.
- Chorpita, B. F., Yim, L., Moffitt, C., Umemoto, L. A., & Francis, S. E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: a revised child anxiety and depression Scale. *Behavior Research and Therapy*, 38, 835-855.
- Costello, E.J., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry*, 60, 837-844.
- de Ross, R., Gullone, E., & Chorpita, B. (2002). The Revised Child Anxiety and Depression Scale: A psychometric investigation with Australian youth. *Behaviour Change*, 19, 90-101.
- DeLong, E., DeLong, D., & Clarke-Pearson, D. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44, 837-845.
- Emons, W., Sijtsma, K., & Meijer, R. (2007). On the Consistency of Individual Classification Using Short Scales. *Psychological Methods*, 12, 105-120.
- Ferdinand, R. (2008). Validity of the CBCL/YSR DSM-IV scales Anxiety Problems and

- Affective Problems. *Journal of Anxiety Disorders*, 22, 126–134.
- Fristad, M. A., Cummins, J., Verducci, J. S., Teare, M., Weller, E. B., & Weller, R. A. (1998a). Study IV: concurrent validity of the DSM-IV revised Children's Interview for Psychiatric Syndromes (ChIPS). *Journal of Child and Adolescent Psychopharmacology*, 8, 227-236.
- Fristad, M. A., Cummins, J., Verducci, J. S., Teare, M., Weller, E. B., & Weller, R. A. (1998b). Study V: Children's Interview for Psychiatric Syndromes (ChIPS) psychometrics in two community samples. *Journal of Child and Adolescent Psychopharmacology*, 8, 237-245.
- Garland, A.F., Kruse, M., Aarons, G.A. (2003). Clinicians and outcome measurement: What's the use? *Journal of Behavioral Health Services and Research*, 30, 393-405.
- Gibbons R.D., Grochocinski, V. J., Weiss, D. J., et al. 2008. Using computerized adaptive testing to reduce the burden of mental health assessment. *Psychological Services*, 59, 361-368.
- Gustafsson, J. & Aberg-Bengtsson, L. (2010). Unidimensionality and the interpretability of psychological instruments. In S. E. Embretson (Ed.), *Measuring Psychological Constructs*. (pp. 97-121). Washington DC: American Psychological Association.
- Gustafsson, J., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, 28, 407–434.
- Harman, H., & Jones, W. (1966) Factor analysis by minimizing residuals (minres). *Psychometrika*, 31, 3, 351-368.
- Holgado-Tello, F., Chacón-Moscoso, S., Barbero-García, I., & Vila-Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality and Quantity*, 44(1), 153-166.
- Holzinger, K. J., & Swineford, R. (1937). The bifactor method. *Psychometrika*, 2, 41-54.

- Hosmer, D., and Lemeshow, S. (2000). 2000. Applied logistic regression, 2nd ed. John Wiley & Sons, Inc., New York.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Kendall, P., Flannery-Schroeder, E., Panichelli-Mindel, S., Southam-Gerow, M., Henin, A., & Warman, M. (1997). Therapy for youths with anxiety disorders: A second randomized clinical trial. *Journal of Consulting and Clinical Psychology*, 65, 366-380.
- Kovacs, M. (1981). Rating scales to assess depression in school-aged children. *Acta Paedopsychiatrica*, 46, 305-315.
- Levitt, J. Saka, N., Romanelli, L. & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology*, 45, 163-191.
- Lewinsohn, P. M., Zinbarg, R., Seeley, J. R., Lewinsohn, M. & Sack, W. H. (1997). Lifetime comorbidity among anxiety disorders and between anxiety disorders and other mental disorders in adolescents. *Journal of Anxiety Disorders*, 11, 377-394.
- McDonald, R. P. (1999). Test theory: A unified approach. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Muris, P. & Meesters, C. (2002). Symptoms of anxiety disorders and teacher-reported school functioning of normal children. *Psychological Report*, 91, 588-590.
- Muris, P., Meesters, C., & Schouten, E. (2002). A brief questionnaire of DSM-IV-defined anxiety and depression symptoms among children. *Clinical Psychology and Psychotherapy*, 9, 430-442.
- Nunnally, J. C. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

- R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>.
- Reise, S., Moore, T., & Haviland, M. (2010). Bifactor Models and Rotations: Exploring the Extent to which Multidimensional Data Yield Univocal Scale Scores. *Journal of Personality Assessment*, 92, 544-559.
- Reise, S., Morizot, J., & Hays, R. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*, 16, 19–31.
- Revelle, W. (2009). psych: Procedures for Psychological, Psychometric, and Personality Research. R package version 1.0-68. <http://personality-project.org/rhttp://personality-project.org/r/psych.manual.pdf>
- Rones, M. & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical Child and Family Psychology Review*, 3, 223-241.
- Satorra, A. & Bentler, P. (1991). Goodness-of-fit test under IV estimation: Asymptotic robustness of a NT test statistic. In R. Gutiérrez & M. J. Valderrama (Eds.), *Applied Stochastic Models and Data Analysis*: pp. 555-567. Singapore: World Scientific.
- Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, 22, 53-61.
- Silverman, W. K., & Albano, A. M. (1996). *Anxiety Disorders Interview Schedule for DSM-IV, Child and Parent Versions*. San Antonio, TX: Psychological Corporation.
- Stark, K., & Laurent, J. (2001). Joint factor analysis of the Children's Depression Inventory and the Revised Children's Manifest Anxiety Scale. *Journal of Clinical Child Psychology*, 30, 552-567.

- Teare, M., Fristad, M. A., Weller, E. B., Weller, R. A., & Salmon, P. (1998a). Study I: development and criterion validity of the Children's Interview for Psychiatric Syndromes (ChIPS). *Journal of Child and Adolescent Psychopharmacology*, 8, 205-211.
- Teare, M., Fristad, M. A., Weller, E. B., Weller, R. A., & Salmon, P. (1998b). Study II: concurrent validity of the DSM-III-R Children's Interview for Psychiatric Syndromes (ChIPS). *Journal of Child and Adolescent Psychopharmacology*, 8, 213-219.
- Van Oort, F.V.A., Greaves-Lord, K., Verhulst, F.C., Ormel, J., & Huizink, A.C. (2009). The developmental course of anxiety symptoms during adolescence: The TRAILS study. *Journal of Child Psychology and Psychiatry*, 50, 1209-1217.
- Weller, E. B., Weller, R. A., Teare, M., Fristad, M. A. (1999). Children's Interview for Psychiatric Syndromes (ChIPS). Washington, DC: American Psychiatric Press.
- Wainer, H., & Thissen, D. (1996). How is reliability related to the quality of test scores? What is the effect of local dependence on reliability? *Educational Measurement: Issues and Practice*, 15(1), 22-29.
- Wolff, H., & Preising, K. (2005). Exploring item and higher order factor structure with the Schmid-Leiman solution: Syntax codes for SPSS and SAS. *Behavioral Research Methods*, 37, 48-58.
- Zinbarg, R. E., Barlow, D. H., & Brown, T. A. (1997). Hierarchical structure and general factor saturation of the anxiety sensitivity index: Evidence and implications. *Psychological Assessment*, 9, 277-284.
- Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α , Revelle's ω , and McDonald's: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70, 123-133.

Table 1

Youth and Caregiver Demographic Information

	Clinic-referred Sample (N=303)		School Sample (N=1060)	
	<i>n</i>	Percentage	<i>n</i>	Percentage
Youth Gender				
Boys	204	67.3	417	39.3
Girls	99	32.7	579	54.6
Missing	--	--	64	6.0
Youth Ethnicity				
Multiethnic	82	27.1	253	23.9
White	147	48.5	16	1.5
African American	32	10.6	17	1.6
Asian American	11	3.6	350	33.0
Latino/Hispanic	22	7.3	2	0.2
Other	9	3.0	19	1.8
Missing	0	0	403	38.1
Caregiver Marital Status				
Married	117	38.6	501	47.3
Divorced, separated	92	30.3	82	7.7
Widowed	14	4.6	9	0.8
Single	51	16.8	53	5.0
Other	21	6.9	2	0.2
Missing	8	2.6	413	39.0
Family Income				
\$0-\$20,000	69	22.8	71	6.7
\$20,001 - \$40,000	99	32.7	109	10.3
\$40,001 - \$60,000	43	14.2	89	8.5
\$60,001 - \$80,000	30	9.9	84	7.9
\$80,001 or more	47	15.5	263	24.9
Missing	15	5.0	444	41.9

Table 2

Diagnoses among the Clinic-referred Youth Sample (N = 303)

Diagnoses	Anywhere	Primary
Depressive Disorders	36	15
Major depressive disorder	24	10
Dysthymic disorder	8	4
Depressive Disorder NOS	4	1
Anxiety Disorders	114	61
Panic disorder and/or agoraphobia	0	0
Specific Phobia	34	18
Generalized anxiety disorder	11	3
Separation anxiety disorder	38	23
Social phobia	15	9
Obsessive-compulsive disorder	10	4
PTSD	4	2
Anxiety NOS	2	2
ADHD	72	33
ADHD-Combined Type	19	4
ADHD-PI	29	13
ADHD-PH	2	4
ADHD-NOS	22	12
Disruptive Behavior Disorders	107	66
Oppositional defiant disorder	73	49
Conduct disorder	33	17
Disruptive behavior disorder NOS	1	0
Schizophrenia	2	1
Bipolar	1	1
Other	6	3
No Diagnosis	123	123
Missing	4	4

Note. **Anywhere** = a diagnosis that appears anywhere in a child's diagnostic profile; **Primary** = a child's primary diagnosis **NOS** = not otherwise specified; **PTSD** = post-traumatic stress disorder; **ADHD** = attention-deficit/hyperactivity disorder; **PI** = Primarily Inattentive Type; **PH** = Primarily Hyperactive Type.